18th International Conference on Information Fusion
Washington, DC - July 6-9, 2015

Statistical Evaluation of Information Source Performance

Dominic E. Schaub
DRDC, Atlantic Research Centre
P.O. Box 1012, 9 Grove Street
Dartmouth, NS, Canada B2Y 3Z7
dominic.schaub@drdc-rddc.gc.ca

***Abstract*—The present work examines the problem of evaluating the performance of statistically-characterized information sources when ground truth is unavailable. Although exact verification may be infeasible, inter-source statistical dependencies may be used to test for information consistency. Through application of a Rosenblatt transformation on an input sample and subsequent Kolmogorov–Smirnov test against the uniform distribution, a given information source can be statistically evaluated for goodness of fit. An algorithm is derived for detecting the presence of suspect information and identifying the associated aberrant source(s). The paper concludes with an example that considers the detection of a malfunctioning radar system in the absence of ground truth.**

***Keywords*—Information Validation, Sensor Performance, Fusion, Kolmogorov–Smirnov Test, Rosenblatt Transformation.**

I. Introduction

Fusion systems depend crucially on the accurate characterizations of information sources in order to properly integrate data in support of the inference being sought. When a collection of data (such as a set of measurements from a physical sensor) does not exhibit its assumed characteristics—because of incorrect model specification or a system malfunction—, the quality of fusion is degraded, leaving downstream decisions vulnerable. It is therefore of considerable importance to rapidly identify aberrant data to minimize subsequent negative outcomes. To this end, information sources may be evaluated directly against ground truth, an approach where the resulting validation (or lack thereof) enjoys a high degree of confidence. However, in many scenarios—particularly in operation centres that receive voluminous data from many diverse and distant assets—this is prohibitively expensive or otherwise infeasible.

One mitigation strategy involves assessing data from multiple distinct sources (such as tracks of a single vessel produced by independent fusion systems) for agreement to within the modeled uncertainty. In this framework, sources are evaluated against their own assumed statistical models and/or compared with other sources that produce redundant data. The manner in which this approach is applied depends significantly on the representation of uncertainty. Existing works have considered sensor validation in a variety of contexts characterized by particular combinations of sensor failure modes, uncertainty representation (Dempster–Shafer, fuzzy, etc), and other domain-specific attributes. In [1], pattern matching and the Nadaraya–Watson estimator were used to identify inconsistent sensor readings in a framework that was applied to pH sensors. Similarly, [2] studied the problem of intermittent and soft-sensor faults in physically redundant sensors, while [3] applied the Kullback–Liebler divergence to verify the functioning of object-localization sensors in the context of microphone selection. In [4], a Bayesian method was developed that identifies inconsistent sensor data through analysis of entropy.

The present work considers a collection of information sources as generalized sensors that form the components of a larger statistical system where inter-source statistical dependencies are assumed be completely described. In this framework, the statistical distributions representing data generated by individual information sources are cast into a multidimensional uniform distribution, allowing samples (data) to be evaluated for goodness of fit. A specialization of this approach is derived for problems that meet certain Markov conditions that commonly accompany target tracking. The special case of the Kalman filter is examined, followed by the development of a more general, non-Gaussian, nonlinear Bayesian smoothing framework. The paper concludes with an example where a group of sensors is used to detect the malfunctioning of a radar system.

II. General Mathematical Framework

The fusion system under consideration is modeled as a statistical system composed of a pair of continuous random vectors Z and X that represent information sources and system state, respectively. It is assumed that Z and X have known distributions and that their interdependence may be fully characterized mathematically. In what follows, information sources are defined as distinct subsets Z^k of Z that are mutually disjoint

$$Z = \left[Z^1, Z^2, \cdots, Z^{|Z|}\right]^\intercal \tag{1}$$

where $|Z|$ denotes the total number of sources in Z. The random vector X may be similarly decomposed into smaller random vectors (which may correspond, for example, to a time-series of states)

$$X = \left[X^1, X^2, \cdots, X^{|X|}\right]^\intercal \quad . \tag{2}$$

Note that in the case of time series, $|Z|$ and $|X|$ are often identical.

A (possibly empty) subset of sources Z_{tr} of Z are assumed to function correctly and provide data consistent with their statistical definitions. These sources are segregated from

their complement Z_{ev}, which form the set of sources to be evaluated. This partition is given as

$$PZ = P\begin{bmatrix} Z^1 \\ \vdots \\ Z^k \\ \vdots \\ Z^{|Z|} \end{bmatrix} = \begin{bmatrix} Z^1_{\mathrm{ev}} \\ \vdots \\ Z^{k'}_{\mathrm{ev}} \\ \vdots \\ Z^{|Z_{\mathrm{ev}}|}_{\mathrm{ev}} \\ Z^0_{\mathrm{tr}} \\ \vdots \\ Z^{k'}_{\mathrm{tr}} \\ \vdots \\ Z^{|Z_{\mathrm{tr}}|}_{\mathrm{tr}} \end{bmatrix} = \begin{bmatrix} Z_{\mathrm{ev}} \\ Z_{\mathrm{tr}} \end{bmatrix} \tag{3}$$

where P is a permutation matrix that reorders Z, and $|Z_{\mathrm{ev}}|$ and $|Z_{\mathrm{tr}}|$ are the total number of measurements for the evaluated and trusted sources, respectively that satisfy $|Z_{\mathrm{ev}}| + |Z_{\mathrm{tr}}| = |Z|$. The permutation matrix maps a given Z^k into either $Z^{k'}_{\mathrm{ev}}$ or $Z^{k'}_{\mathrm{tr}}$ for some k' where generally $k \neq k'$. Finally, each Z^k_{ev} may be further decomposed into its individual scalar components

$$Z^k_{\mathrm{ev}} = \left(Z^{k,1}_{\mathrm{ev}}, Z^{k,2}_{\mathrm{ev}}, \ldots, Z^{k,|Z^k_{\mathrm{ev}}|}_{\mathrm{ev}}\right)^{\intercal} \tag{4}$$

where $|Z^k_{\mathrm{ev}}|$ denotes the number of elements in the k^{th} evaluated information source, and the total number of degrees of freedom of Z_{ev} is given by

$$n = \sum_{k=1}^{|Z_{\mathrm{ev}}|} |Z^k_{\mathrm{ev}}| \ . \tag{5}$$

Both Z and X completely describe the system over the entire period of time under consideration. In systems characterized by successive states and/or measurements, the indexed members of the associated times series are mapped to distinct Z^k and X^k. Consequently, a statistical sample representing the data produced by the information sources will comprise only a single point in the high-dimensional space spanned by Z.[1] The approach adopted in the present work evaluates a one-point sample $\mathfrak{Z}$ of Z for statistical consistency against the model given by (Z, X). Although technically redundant, specification of X at the outset of analysis may simplify the conceptualization of the problem, which typically involves estimating some component(s) of the system state. The range of the doublet (Z, X) is given by the space $\mathbb{S}$ that is formed by the Cartesian product

$$\mathbb{S} = \mathbb{S}_Z \times \mathbb{S}_X \tag{6}$$

where $\mathbb{S}_Z$ and $\mathbb{S}_X$ are spanned by the members of Z and X, respectively.

The fusion system derives its utility from the statistical dependencies between the various Z^k and X^l. By design, these dependencies (which may be represented by copulas) relate information and its uncertainty with the statistical inference that is sought. For example, in the context of target tracking, the elements of Z are sensor measurements while those of X represent the system state (e.g. target position and velocity) from which estimates may be computed. The members of Z and X, which are indexed by time, exhibit dependencies resulting from the prescribed sensor likelihood functions (that model measurement processes) and transition densities (that describe target motion such as that due to Newtonian mechanics). Furthermore, Markov conditions are usually imposed, greatly simplifying the structure of (Z, X), thereby allowing successive state vectors to be computed recursively as new measurements become available.

In the present setting, it is assumed possible to define the system joint probability density function for (Z, X) over $\mathbb{S}$ denoted by

$$f_{Z,X}\left(\mathbf{z}, \mathbf{x}\right) \tag{7}$$

where $\mathbf{z}$ and $\mathbf{x}$ are the vectors that may be decomposed as subvectors

$$\mathbf{z} = \left[\mathbf{z}^1, \ldots, \mathbf{z}^{|Z|}\right]^{\intercal}, \quad \text{and} \quad \mathbf{x} = \left[\mathbf{x}^1, \ldots, \mathbf{x}^{|X|}\right]^{\intercal} \ . \tag{8}$$

Upon definition of (7), members of X are no longer required and may be regarded as nuisance variables that can be eliminated through marginalization

$$f_Z\left(\mathbf{z}\right) = \int_{\mathbb{S}_X} f_{Z,X}\left(\mathbf{z}, \mathbf{x}\right) d\mathbf{x} \ . \tag{9}$$

The structure of the information source evaluation problem is captured by $f_Z(\mathbf{z})$, which defines the statistical relationships between the various sources.

Analysis proceeds by incorporating data from the trusted information sources in the form of the sample

$$\mathfrak{Z}_{\mathrm{tr}} = (\mathfrak{z}^1_{\mathrm{tr}}, \mathfrak{z}^2_{\mathrm{tr}}, \ldots, \mathfrak{z}^{|Z_{\mathrm{tr}}|}_{\mathrm{tr}}) \tag{10}$$

which yields the conditional density function

$$f_{Z_{\mathrm{ev}}|Z_{\mathrm{tr}}}\left(\mathbf{z}_{\mathrm{ev}}|\mathbf{z}_{\mathrm{tr}} = \mathfrak{Z}_{\mathrm{tr}}\right) \propto f_Z\left(\mathbf{z}_{\mathrm{ev}}, \mathbf{z}_{\mathrm{tr}} = \mathfrak{Z}_{\mathrm{tr}}\right) \ . \tag{11}$$

The associated random variable $(Z_{\mathrm{ev}}|Z_{\mathrm{tr}} = \mathfrak{Z}_{\mathrm{tr}})$ incorporates everything that is known about the problem, and thus provides the best statistical model against which a sample from Z_{ev} may be evaluated. Data from the information source under evaluation,

$$\mathfrak{Z}_{\mathrm{ev}} = \left(\mathfrak{z}^1_{\mathrm{ev}}, \mathfrak{z}^2_{\mathrm{ev}}, \ldots, \mathfrak{z}^{|Z_{\mathrm{ev}}|}_{\mathrm{ev}}\right) \tag{12}$$

may thus be carried out by examining its goodness of fit against (11). This process may be formalized by first transforming $(Z_{\mathrm{ev}}|Z_{\mathrm{tr}} = \mathfrak{Z}_{\mathrm{tr}})$ into the n-dimensional uniform distribution $Y = (Y^1, Y^2, \ldots, Y^{|Z_{\mathrm{ev}}|})$ over $[0,1]^{\prod_k |Z^k_{\mathrm{ev}}|}$ by way of the Rosenblatt transformation (Appendix). The transformed probability density function becomes

$$f_Y\left(\mathbf{y}\right) = \prod_{k=1}^{|Z_{\mathrm{ev}}|} \prod_{i=1}^{|Z^k_{\mathrm{ev}}|} \left(H\left(y^{k,i}\right) - H\left(y^{k,i} - 1\right)\right) \tag{13}$$

where $H(\cdot)$ is the Heaviside step function. A given sample $\mathfrak{Z}_{\mathrm{ev}}$ of Z is transformed into the sample $\mathfrak{Y}_{\mathrm{ev}}$ of Y using

$$\mathfrak{Y}_{\mathrm{ev}} = R\left(\mathfrak{Z}_{\mathrm{ev}}\right) \tag{14}$$

[1] For example, an n-element i.i.d. sample drawn from a scalar distribution given by $Y \sim \mathcal{Y}$ may be regarded equivalently as a one-point sample from the vector $(Y_1, \ldots, Y_n)$ where $Y_1 \sim \mathcal{Y}, \ldots, Y_n \sim \mathcal{Y}$. The utility of a one point sample in statistical inference thus depends crucially on the dependency structure between elements of the presumed random vector. Further discussion may be found in §III.

where $R(\cdot)$ is given by (39) in Definition VII.1. When the evaluated sources are properly modeled, the transformed sample $\mathfrak{Y} = \{\mathfrak{y}^1, \mathfrak{y}^2, \ldots, \mathfrak{y}^{|Z_{\mathrm{ev}}|}\}$ comprises a set of n independent and identically distributed (i.i.d.) subcomponents $\mathfrak{y}^{k,i}$ drawn from $\mathcal{U}(0,1)$.

Finally, it is important to note that finding the Rosenblatt transform entails computing the series of $|Z_{\mathrm{ev}}|$ conditional cumulative distribution functions

$$\begin{aligned} &F_{Z_{\mathrm{ev}}^{1,1}|Z_{\mathrm{tr}}}\left(\mathbf{z}_{\mathrm{ev}}^{1,1} = \mathfrak{z}_{\mathrm{ev}}^{1,1} | \mathbf{z}_{\mathrm{tr}} = \mathfrak{Z}_{\mathrm{tr}}\right) \\ &F_{Z_{\mathrm{ev}}^{\phi_1}|Z_{\mathrm{ev}}^{1,1},Z_{\mathrm{tr}}}\left(\mathbf{z}_{\mathrm{ev}}^{\phi_1} = \mathfrak{z}_{\mathrm{ev}}^{\phi_1} | \mathbf{z}_{\mathrm{ev}}^{1,1} = \mathfrak{z}_{\mathrm{ev}}^{1,1}, \mathbf{z}_{\mathrm{tr}} = \mathfrak{Z}_{\mathrm{tr}}\right) \\ &\qquad\qquad\vdots \\ &F_{Z_{\mathrm{ev}}^{\phi_2}|Z_{\mathrm{ev}}^{\phi_3},Z_{\mathrm{tr}}}\left(\mathbf{z}_{\mathrm{ev}}^{\phi_2} = \mathfrak{z}_{\mathrm{ev}}^{\phi_2} | \mathbf{z}_{\mathrm{ev}}^{(1,1):\phi_3} = \mathfrak{z}_{\mathrm{ev}}^{(1,1):\phi_3}, \mathbf{z}_{\mathrm{tr}} = \mathfrak{Z}_{\mathrm{tr}}\right) \end{aligned} \tag{15}$$

where

$$\begin{aligned} \phi_1 &= \left\{ \begin{array}{ll} (2,1), & |Z_1| = 1 \\ (1,2), & |Z_1| > 1 \end{array} \right\} \\ \phi_2 &= (|Z_{\mathrm{ev}}|, |Z_{\mathrm{ev}}^{|Z_{\mathrm{ev}}|}|) \\ \phi_3 &= \left\{ \begin{array}{ll} (|Z_{\mathrm{ev}}| - 1, |Z_{\mathrm{ev}}^{|Z_{\mathrm{ev}}|-1}|), & \left|Z_{\mathrm{ev}}^{|Z_{\mathrm{ev}}|}\right| = 1 \\ (|Z_{\mathrm{ev}}|, |Z_{\mathrm{ev}}^{|Z_{\mathrm{ev}}|}| - 1), & \left|Z_{\mathrm{ev}}^{|Z_{\mathrm{ev}}|}\right| > 1 \end{array} \right\}. \end{aligned} \tag{16}$$

As there may exist several means to evaluate the F in (15), it is preferable to select an approach that minimizes computation time.

III. Information Source Evaluation

A. Statistical Hypothesis Testing

Evaluation of a sample $\mathfrak{Z}_{\mathrm{ev}}$ against the model given by the random variable $(Z_{\mathrm{ev}}|Z_{\mathrm{tr}} = \mathfrak{Z}_{\mathrm{tr}})$ commences by defining a collection of mutually-exclusive hypotheses. In the present work these are

H_0: The subcomponents of $Y^{k,i}$ of $(Y^1, \ldots, Y^{|Z_{\mathrm{ev}}|})$ are i.i.d., and $Y^{k,i} \sim \mathcal{U}(0,1)$. This is the null hypothesis that asserts that the (Z, X) model is correct.

H_{A}: This is the collection of alternate hypotheses.

An elementary statistical hypothesis test amounts to deciding whether there is adequate evidence to reject H_0 [5], [6]. This involves computing the sample's test statistic D and determining the critical region R_α (not necessarily contiguous) for a predefined significance level α. Under the null hypothesis, D evaluates to within R_α with a probability α. For a sufficiently small significance level, such an event is very unlikely and its occurrence is therefore grounds on which the null hypothesis may be rejected. In many cases (such as one-tailed tests), the critical region may be completely specified by a minimum value K_α such that $R_\alpha = \{x | x \in \mathbb{R} \wedge x > K_\alpha\}$.

Failures to reject the null hypothesis in circumstances where a member of H_A holds (or vice verse) are regarded as errors that may be categorized as

Type I: The null hypothesis H_0 was rejected when in fact it is correct.

Type II: The null hypothesis was not rejected when in fact it is incorrect.

The frequencies of committing type I and type II errors are quantified inversely by the test's specificity and sensitivity (or power), respectively. In general, different tests will exhibit varying error profiles, and moreover, may differ markedly in sensitivity to certain members of H_{A}. Where a particular set of alternate hypotheses is anticipated, it may be desirable (when possible) to tailor tests to maximize sensitivity to this group.

The Kolmogorov–Smirnov test (Appendix) may be employed as a general technique for evaluating i.i.d. samples for goodness of fit with respect to a predefined cumulative distribution function. In the present work, it was used to test the subcomponents of $\mathfrak{Y}_{\mathrm{ev}}$ against the uniform distribution $\mathcal{U}(0,1)$. In each case, a significance level is chosen, and the test statistic is computed and compared against the associated critical value, which is found in a reference table or by numerical computation.

In most cases, analysis of type II errors is challenging and requires the definition of alternative hypotheses. In certain applications these may be readily identified, and the Kolmogorov–Smirnov test can be modified to improve sensitivity [7], [8]. Qualitatively, however, the test's sensitivity depends crucially on the mathematical structure of the assumed model. Where the random variables $Z_{\mathrm{ev}}^{k'}$ in $(Z_{\mathrm{ev}}|Z_{\mathrm{tr}} = \mathfrak{Z}_{\mathrm{tr}})$ are i.i.d., the problem becomes essentially one dimensional, and the Rosenblatt transformation is found trivially. In this instance, the alternative hypotheses will comprise random variables whose associated joint density function can be factored into independent terms, and the Kolmogorov–Smirnov test will exhibit the power typically expected with a univariate i.i.d. sample. However, should the statistical structure of $(Z_{\mathrm{ev}}|Z_{\mathrm{tr}} = \mathfrak{Z}_{\mathrm{tr}})$ be complex, the problem's dimensionality may be higher, and the test becomes significantly less powerful. In cases characterized by moderate degrees of statistical dependence, the test will likely be of intermediate sensitivity.

Conditioning on Z_{tr} improves the performance of hypothesis testing by refining $f_Z(\mathbf{z})$ to more accurately reflect the given circumstances, thereby reducing the incidence of type II errors (the type-I error rate should remain unaffected because it is explicitly defined). Conditioning can also be viewed in terms of mutual information between Z_{tr} and Z_{ev} [9]. In cases where the mutual information is high, knowledge of Z_{tr} substantially reduces the uncertainty in Z_{ev}. Qualitatively, this is expected to improve the test's sensitivity.

B. Kolmogorov–Smirnov Statistic Under Source Exclusion

The foregoing hypothesis test is appropriate when certain sources are believed to produce tainted data. In many situations however, the suspect source(s) may not be obvious, and an alternative approach is required. In these cases, the absence of readily-identified trusted sources impedes the straightforward application of hypothesis testing. One possible (but problematic) solution is based on a partition of Z that designates a single Z^k as a evaluated and the remaining $Z \setminus Z^k$ as trusted. The set of all such partitions may then be used

to carry out a sequence of hypothesis tests. Unfortunately, the resultant test statistics (which are not independent) will incur significant error by the inclusion of the problematic source, and thus $(Z_{\text{ev}}|Z_{\text{tr}} = \mathfrak{Z}_{\text{tr}})$ may deviate substantially from the correct model. An alternative approach is based on sequentially *excluding* a single Z^k at a time, and computing the Kolmogorov–Smirnov statistic on the remaining $|Z| - 1$ sources. In the calculation performed without the problematic source, the statistic is not expected to exceed the critical value. In all other cases—depending on the test's sensitivity—the computed statistics may deviate noticeably from what is expected under H_0. Although the resultant D are once again not statistically independent, this may not be of serious concern in practical applications characterized by large numbers of data.

IV. Efficient Joint PDF Reductions for Markov Transitions and Measurements

The dimensionality of (11) may be prohibitive in practical settings. However, when applied to recursive Bayesian filtering [10]–[12], the mathematical framework of the foregoing section admits a particularly simple decomposition. In this setting, a time-series pair of random vectors X^k and Z^k represent system states and observations, respectively. In accordance with (3), the Z^k are embedded as subvectors in Z, and X is decomposed analogously. A prior probability density function $f_{X^1}(\mathbf{x}^1)$ is assigned over a space of possible system states $\mathbb{S}_X$. This function is then alternately brought forward in time using a transition function $f_{X^{k+1}|X^k}(\mathbf{x}^{k+1}|\mathbf{x}^k)$ that describes the state evolution

$$f_{X^{k+1}|Z^{1:k}}\left(\mathbf{x}^{k+1}|\mathbf{z}^{1:k}\right) = \int_{\mathbb{S}_X} f_{X^{k+1}|X^k}\left(\mathbf{x}^{k+1}|\mathbf{x}^k\right) f_{X^k|Z^{1:k}}\left(\mathbf{x}^k|\mathbf{z}^{1:k}\right) d\mathbf{x}^k \tag{17}$$

and then refined with a Bayesian likelihood function $f_{Z^k|X^k}(\mathbf{z}^k|\mathbf{x}^k)$ that represents a system observation

$$f_{X^k|Z^{1:k}}\left(\mathbf{x}^k|\mathbf{z}^{1:k}\right) = \frac{f_{Z^k|X^k}\left(\mathbf{z}^k|\mathbf{x}^k\right) f_{X^k|Z^{1:k-1}}\left(\mathbf{x}^k|\mathbf{z}^{1:k-1}\right)}{\displaystyle\int_{\mathbb{S}_X} f_{Z^k|X^k}\left(\mathbf{z}^k|\mathbf{x}^k\right) f_{X^k|Z^{1:k-1}}\left(\mathbf{x}^k|\mathbf{z}^{1:k-1}\right) d\mathbf{x}^k} \tag{18}$$

where $\mathbf{z}^{1:k}$ and $\mathbf{z}^{1:k-1}$ are subvectors of $\mathbf{z}$ given in (8). Should multiple observations occur simultaneously the (zero-time) transition between timesteps is omitted. In practice, the $\mathbf{z}^k$ are replaced with actual observations $\mathfrak{Z}^k$.

In many filtering applications, Markov conditions are imposed on both the transition and measurement functions, yielding, respectively,

$$f_{X^{k+1}|X^k}\left(\mathbf{x}^{k+1}|\mathbf{x}^k\right) = f_{X^{k+1}|X^{1:k}}\left(\mathbf{x}^{k+1}|\mathbf{x}^{1:k}\right) \tag{19}$$

and

$$f_{Z^{k+1}|X^{k+1}}\left(\mathbf{z}^{k+1}|\mathbf{x}^{k+1}\right) = f_{Z^{k+1}|X^{1:k+1}}\left(\mathbf{z}^{k+1}|\mathbf{x}^{1:k+1}\right) \tag{20}$$

where $\mathbf{x}^{1:k}$ and $\mathbf{x}^{1:k+1}$ are subvectors of $\mathbf{x}$ defined in (8). These simplifications enable the recursive formulation of (17) and (18) and render tractable a problem that would otherwise be of prohibitively high dimensionality.

Conditions (19) and (20) ensure that predictions and measurements depend only on the present state, and thus the effective problem dimension reduces to that of a single time-indexed state vector. Furthermore, under the Markov constraints, the state of a system comprising $|Z|$ observations is given by a joint probability density function that may be factored as [10]

$$f_{X^{1:|Z|},Z^{1:|Z|}}\left(\mathbf{x}^{1:|Z|},\mathbf{z}^{1:|Z|}\right) = f_{X^1}\left(\mathbf{x}^1\right)\prod_{k=2}^{|Z|} f_{Z^k|X^k}\left(\mathbf{z}^k|\mathbf{x}^k\right) f_{X^k|X^{k-1}}\left(\mathbf{x}^k|\mathbf{x}^{k-1}\right) \ . \tag{21}$$

Equation (21) serves as the joint density in (7) and may be marginalized and conditioned on $\mathfrak{Z}^k_{\text{tr}}$ to yield

$$f_{Z_{\text{ev}}}\left(\mathbf{z}_{\text{ev}}\right) \propto \int_{\mathbb{S}_X} f_{X^1}\left(\mathbf{x}^1\right)\prod_{k=2}^{|Z|} f_{Z^k|X^k}\left(\mathbf{z}^k_{\text{ev}},\mathbf{z}^k_{\text{tr}}=\mathfrak{Z}^k_{\text{tr}}|\mathbf{x}^k\right) f_{X^k|X^{k-1}}\left(\mathbf{x}^k|\mathbf{x}^{k-1}\right) d\mathbf{x} \ . \tag{22}$$

The method employed to compute the conditional functions given in (15) depends on the problem's latent mathematical structure (and the structure of $f_{Z_{\text{ev}}}(\mathbf{z}_{\text{ev}})$ in particular). Where each term in (22) is distributed normally, $f_{Z_{\text{ev}}}(\mathbf{z}_{\text{ev}})$ may be found algebraically. Otherwise, smoothing techniques must be employed. These methods are described below.

A. Linear Gaussian Case

In this case, the transition density and measurement likelihood are given by [11]

$$f_{X^{k+1}|X^k}\left(\mathbf{x}^{k+1}|\mathbf{x}^k\right) \propto e^{-\frac{1}{2}\left(\mathbf{x}^{k+1}-F_{X^{k+1}|X^k}\mathbf{x}^k\right)^{\intercal}\Sigma^{-1}_{X^{k+1}|X^k}\left(\mathbf{x}^{k+1}-F_{X^{k+1}|X^k}\mathbf{x}^k\right)} \tag{23}$$

and

$$f_{Z^k|X^k}\left(\mathbf{z}^k|\mathbf{x}^k\right) \propto e^{-\frac{1}{2}\left(\mathbf{z}^k-H_{Z^k|X^k}\mathbf{x}^k\right)^{\intercal}\Sigma^{-1}_{Z^k|X^k}\left(\mathbf{z}^k-H_{Z^k|X^k}\mathbf{x}^k\right)} \tag{24}$$

where $F_{X^{k+1}|X^k}$ and $H_{Z^k|X^k}$ are matrices that describe target kinematics and the measurement process, respectively. The proportionality relations in (23) and (24) may be replaced by equalities by scaling the respective right-hand sides with factors related to the covariance-matrix determinants (although this is unnecessary here). Equations (23) and (24) may be substituted into (21) to yield a single, consolidated multivariate Gaussian function

$$f_{Z^{1:|Z|},X^{1:|Z|}}\left(\mathbf{z}^{1:|Z|}|\mathbf{x}^{1:|Z|}\right) \propto e^{-\frac{1}{2}\left(\begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix}-\begin{bmatrix}\mu_{\mathbf{x}}\\ \mu_{\mathbf{z}}\end{bmatrix}\right)^{\intercal}\Sigma^{-1}\left(\begin{bmatrix}\mathbf{x}\\ \mathbf{z}\end{bmatrix}-\begin{bmatrix}\mu_{\mathbf{x}}\\ \mu_{\mathbf{z}}\end{bmatrix}\right)} \tag{25}$$

where $\mathbf{z}$ and $\mathbf{x}$ are notational abbreviations of $\mathbf{z}^{1:|Z|}$ and $\mathbf{x}^{1:|Z|}$, respectively. The system covariance matrix Σ and vector means $\mu_{\mathbf{z}}$ and $\mu_{\mathbf{x}}$ are easily constructed from the prior and individual term-term relations defined in (23) and (24). Marginalization of $\mathbf{x}$ requires factorization of Σ^{-1} in a manner that separates the elements of $\mathbf{z}$ from those of $\mathbf{x}$. Noting that the covariance matrix is symmetric, its inverse may be decomposed as

$$\Sigma^{-1} = \begin{bmatrix} A & B \\ B^{\intercal} & C \end{bmatrix} \tag{26}$$

where A, B, $B^\intercal$, and C correspond to the $\mathbf{xx}$, $\mathbf{xz}$, $\mathbf{zx}$, and $\mathbf{zz}$ terms, respectively. In this form, Σ^{-1} may be factored by the Aitken block diagonalization [13], [14]

$$\Sigma^{-1} = \begin{bmatrix} I & 0 \\ B^\intercal A^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & \Sigma^{-1}/A \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix} \tag{27}$$

where $\Sigma^{-1}/A = C - B^\intercal A^{-1} B$ is the Shur complement. Equation (25) thus simplifies to

$$f_{Z^{1:|Z|},X^{1:|Z|}}\left(\mathbf{z}^{1:|Z|},\mathbf{x}^{1:|Z|}\right) \propto$$
$$e^{-\frac{1}{2}\left(\mathbf{x}-\mu_\mathbf{x}-A^{-1}B\left(\mathbf{z}+\mu_\mathbf{z}\right)\right)^\intercal A\left(\mathbf{x}-\mu_\mathbf{x}-A^{-1}B\left(\mathbf{z}+\mu_\mathbf{z}\right)\right)}$$
$$\cdot e^{-\frac{1}{2}\left(\mathbf{z}-\mu_\mathbf{z}\right)^\intercal\left[\Sigma^{-1}/A\right]\left(\mathbf{z}-\mu_\mathbf{z}\right)}. \tag{28}$$

Noting that the term $\mu_\mathbf{x} - A^{-1}B(\mathbf{z}+\mu_\mathbf{z})$ may be treated as a mean, marginalization in accordance with (9) yields

$$f_{Z^{1:|Z|}}\left(\mathbf{z}^{1:|Z|}\right) \propto e^{-\frac{1}{2}\left(\mathbf{z}-\mu_\mathbf{z}\right)^\intercal \Sigma_\mathbf{z}^{-1}\left(\mathbf{z}-\mu_\mathbf{z}\right)} \tag{29}$$

where $\Sigma_\mathbf{z}^{-1} = \Sigma^{-1}/A$.

Conditioning on $\mathfrak{Z}_\mathrm{tr}$ may be carried out by further separating $\mathbf{z}_\mathrm{tr}$ from $\mathbf{z}_\mathrm{ev}$ through a decomposition of $\Sigma_\mathbf{z}^{-1}$ as

$$\Sigma_\mathbf{z}^{-1} = \begin{bmatrix} D & E \\ E^\intercal & F \end{bmatrix} \tag{30}$$

where D, E, $E^\intercal$, and F correspond to the $\mathbf{z}_\mathrm{ev}\mathbf{z}_\mathrm{ev}$, $\mathbf{z}_\mathrm{ev}\mathbf{z}_\mathrm{tr}$, $\mathbf{z}_\mathrm{tr}\mathbf{z}_\mathrm{ev}$, and $\mathbf{z}_\mathrm{tr}\mathbf{z}_\mathrm{tr}$ terms, respectively. A second Aitken block diagonalization yields

$$f_{Z_\mathrm{ev}}\left(\mathbf{z}_\mathrm{ev}\right) \propto e^{-\frac{1}{2}\left(\mathbf{z}_\mathrm{ev}-\mu'_\mathbf{z}\right)^\intercal \Sigma_\mathbf{z}^{-1}\left(\mathbf{z}_\mathrm{ev}-\mu'_\mathbf{z}\right)} \tag{31}$$

where $\mu'_\mathbf{z} = \mu_\mathbf{z} - (D^{-1}E)\mathfrak{Z}_\mathrm{tr}$. The series of conditional functions given in (15) may now be found trivially.

B. General Case

A more general approach must be developed for problems characterized by non-Gaussian and/or nonlinear transition densities and measurement functions. Provided that the Markov conditions of (19) and (20) still hold, the challenges posed by joint density function's dimensionality may be addressed through Bayesian smoothing. The following adapts two smoothers (forward-backward and two-filter) to the source evaluation problem. Both approaches seek to compute the series of conditional density functions

$$\begin{aligned} &f_{Z_\mathrm{ev}^{1,1}|Z_\mathrm{tr}}\left(\mathbf{z}_\mathrm{ev}^{1,1}=\mathfrak{z}_\mathrm{ev}^{1,1}|\mathbf{z}_\mathrm{tr}=\mathfrak{Z}_\mathrm{tr}\right) \\ &f_{Z_\mathrm{ev}^{\phi_1}|Z_\mathrm{tr},Z_\mathrm{ev}^{1,1}}\left(\mathbf{z}_\mathrm{ev}^{\phi_1}=\mathfrak{z}_\mathrm{ev}^{\phi_1}|\mathbf{z}_\mathrm{ev}^{1,1}=\mathfrak{z}_\mathrm{ev}^{1,1},\mathbf{z}_\mathrm{tr}=\mathfrak{Z}_\mathrm{tr}\right) \\ &\qquad\qquad\vdots \\ &f_{Z_\mathrm{ev}^{\phi_2}|Z_\mathrm{tr},Z_\mathrm{ev}^{\phi_3}}\left(\mathbf{z}_\mathrm{ev}^{\phi_2}=\mathfrak{z}_\mathrm{ev}^{\phi_2}|\mathbf{z}_\mathrm{ev}^{(1,1):\phi_3}=\mathfrak{z}_\mathrm{ev}^{(1,1):\phi_3},\mathbf{z}_\mathrm{tr}=\mathfrak{Z}_\mathrm{tr}\right) \end{aligned} \tag{32}$$

where the various ϕ are given in (16). These functions may then be used to find the cumulative distribution functions of (15). The individual observations will be of the form $f_{Z^k|X^k}(\mathbf{z}^k = \mathfrak{z}^k|\mathbf{x}^k) = f_{Z_\mathrm{ev}^{k'}|X^k}(\mathbf{z}_\mathrm{ev}^{k'} = \mathfrak{z}_\mathrm{ev}^{k'}|\mathbf{x}^k)$ where k is related to k' by the permutation matrix of (3). Each smoother solves for $f_{X^k|Z_\mathrm{ev}^{1:k'-1},Z_\mathrm{tr}}(\mathbf{x}^k|\mathbf{z}_\mathrm{ev}^{1:k'-1} = \mathfrak{z}_\mathrm{ev}^{1:k'-1}, \mathbf{z}_\mathrm{tr} = \mathfrak{Z}_\mathrm{tr})$, which, by the law of total probability yields

$$\begin{aligned} &f_{Z_\mathrm{ev}^{k'}|Z_\mathrm{ev}^{1:k'-1},Z_\mathrm{tr}}\left(\mathbf{z}_\mathrm{ev}^{k'}=\mathfrak{z}_\mathrm{ev}^{k'}|\mathbf{z}_\mathrm{ev}^{1:k'-1}=\mathfrak{z}_\mathrm{ev}^{1:k'-1},\mathbf{z}_\mathrm{tr}=\mathfrak{Z}_\mathrm{tr}\right) \\ &=\int f_{Z_\mathrm{ev}^{k'}|X^k}\left(\mathbf{z}_\mathrm{ev}^{k'}=\mathfrak{z}_\mathrm{ev}^{k'}|\mathbf{x}^k\right) \\ &\quad\cdot f_{X^k|Z_\mathrm{ev}^{1:k'-1},Z_\mathrm{tr}}\left(\mathbf{x}^k|\mathbf{z}_\mathrm{ev}^{1:k'-1}=\mathfrak{z}_\mathrm{ev}^{1:k'-1},\mathbf{z}_\mathrm{tr}=\mathfrak{Z}_\mathrm{tr}\right)d\mathbf{x}^k \end{aligned} \tag{33}$$

The left-hand side of (33) is the joint density function of the scalar components $Z_\mathrm{ev}^{k',i}$ of the k'^{th} information source, conditioned on all $k'' < k'$. This function may be further sequentially conditioned for each $Z_\mathrm{ev}^{k',i}$ with $i \in \{1,\ldots,|Z_\mathrm{ev}^{k'}|\}$ to yield the terms of (32). Note that this step depends on the specific dependence structure[2] between the $Z_\mathrm{ev}^{k',i}$.

1) Forward-Backward Smoothing: The standard forward-backward smoother pairs the ordinary forward filtering problem with a second recursion carried out in reverse. For any given X^k, the latter step may be carried out with the pair [15], [16]

$$\begin{aligned} &f_{X^{k+1}|Z^{1:k}}\left(\mathbf{x}^{k+1}|\mathbf{z}^{1:k}\right) = \\ &\int f_{X^{k+1}|X^k}\left(\mathbf{x}^{k+1}|\mathbf{x}^k\right) f_{X^k|Z^{1:k}}\left(\mathbf{x}^k|\mathbf{z}^{1:k}\right)d\mathbf{x}^k \end{aligned} \tag{34}$$

and

$$\begin{aligned} &f_{X^k|Z^{1:|Z|}}\left(\mathbf{x}^k|\mathbf{z}^{1:|Z|}\right) = f_{X^k|Z^{1:k}}\left(\mathbf{x}^k|\mathbf{z}^{1:k}\right) \\ &\cdot\int \frac{f_{X^{k+1}|X^k}\left(\mathbf{x}^{k+1}|\mathbf{x}^k\right) f_{X^{k+1}|Z^{1:|Z|}}\left(\mathbf{x}^{k+1}|\mathbf{z}^{1:|Z|}\right)}{f_{X^{k+1}|Z^{1:|Z|}}\left(\mathbf{x}^{k+1}|\mathbf{z}^{1:|Z|}\right)}d\mathbf{x}^{k+1}. \end{aligned} \tag{35}$$

Note that (34) is simply the Markov transition step that occurs in the forward filter. The backward iteration begins with $X^{|Z|}$ and recursively computes $X^{|Z|-1}, X^{|Z|-2}, \ldots, X^k$. In general, calculation of each successive member in (32) requires performing both a forward and reverse pass. However, a significant portion of each forward filter run may be reused by adopting a reversed conditioning sequence given the by subscripts[3] $(|Z_\mathrm{ev}|, |Z_\mathrm{ev}^{|Z_\mathrm{ev}|}|), (|Z_\mathrm{ev}|, |Z_\mathrm{ev}^{|Z_\mathrm{ev}|}| - 1), \ldots (1,1)$.

2) Two-Filter Smoothing: The two-filter smoother is based on the decomposition [15]

$$\begin{aligned} f_{X^k|Z^{1:|Z|}}\left(\mathbf{x}^k|\mathbf{z}^{1:|Z|}\right) &\propto f_{X^k|Z^{1:k}}\left(\mathbf{x}^k|\mathbf{z}^{1:k}\right) \\ &\cdot f_{Z^{k+1:|Z|}|X^k}\left(\mathbf{z}^{k+1:|Z|}|\mathbf{x}^k\right) \end{aligned} \tag{36}$$

where the leading term on the right-hand side is the probability density function of the state conditioned on all past measurements (and is thus the output of the ordinary forward filter).

[2] This step is trivial when the $Z_\mathrm{ev}^{k',1},\ldots,Z_\mathrm{ev}^{k',|Z_\mathrm{ev}^{k'}|}$ are statistically independent.

[3] The second term of this sequence is $(|Z_\mathrm{ev}| - 1, |Z_\mathrm{ev}^{|Z_\mathrm{ev}-1|}|)$ when $|Z_\mathrm{ev}^{|Z_\mathrm{ev}|}| = 1$.

Applying Bayes' rule, the trailing term becomes

$$f_{Z^{k+1:|Z|}|X^k}\left(\mathbf{z}^{k+1:|Z|}|\mathbf{x}^k\right) \propto \frac{f_{X^k|Z^{k+1:|Z|}}\left(\mathbf{x}^k|\mathbf{z}^{k+1:|Z|}\right)}{f_{X^k}\left(\mathbf{x}^k\right)} \tag{37}$$

when $\mathbf{z}^{k+1:|Z|}$ is fixed. The numerator on the right-hand side is produced by a filter running backwards in time (analogous in function to $f_{X^k|Z^{1:k}}(\mathbf{x}^k|\mathbf{z}^{1:k})$ from the forward filter), and the denominator is related to $f_{X^{|Z|}}(\mathbf{x}^{|Z|})$ by a series of reverse Markov transitions. The two-filter smoother is adapted to computing the conditional densities of (33) by running the forward filter with only $Z_{\text{tr}}^{\leq k} = \mathfrak{Z}_{\text{tr}}^{\leq k}$ and the reverse filter with both $Z_{\text{tr}}^{>k} = \mathfrak{Z}_{\text{tr}}^{>k}$ and $Z_{\text{ev}}^{>k} = \mathfrak{Z}_{\text{ev}}^{>k}$ (i.e. $Z^{>k} = \mathfrak{Z}^{>k}$). The joint density functions for each $Z_{\text{ev}}^{k'} = \mathfrak{Z}_{\text{ev}}^{k'}$ are found in reverse, starting with $Z_{\text{ev}}^{|Z_{\text{ev}}|} = \mathfrak{Z}_{\text{ev}}^{|Z_{\text{ev}}|}$. Note that this approach requires each of the forward and reverse runs to be performed only once.

Although this smoother may enjoy a straightforward implementation (such as in the linear-Gaussian case), the final probability density $f_{X^{|Z|}}(\mathbf{x}^{|Z|})$ in the reverse filter must be sufficiently uninformative so as to ensure compatibility with (37). In practice this entails using a $f_{X^{|Z|}}(\mathbf{x}^{|Z|})$ of very high variance, which may engender numerical stability and tractability issues. Further detail may be found in [15].

C. Computational Complexity

The running time of the reduction algorithms will depend on a number of parameters. The asymptotic computational complexity as a function of the total number of information sources (or timesteps), however, is the most significant characteristic that determines the scalability in most applications. The complexity of the matrix-based linear Gaussian method (§IV-A) depends primarily on the cost of matrix inversion incurred by calculating the shifts in $\mu_{\mathbf{z}}$, whose naïve complexity is $\mathcal{O}(|Z_{\text{ev}}|^3)$. However, as the constituent (non-inverted) submatrices are sparse, a carefully-designed algorithm may yield significant improvements in running time.

The forward-backward algorithm of (IV-B1) requires both a forward and backward pass for each evaluated source. As a single pass entails $\mathcal{O}(|Z_{\text{ev}}|)$ numerical operations, the complexity of the entire problem is quadratic in $|Z_{\text{ev}}|$ (i.e. $\mathcal{O}(|Z_{\text{ev}}|^2)$). In contrast, the two-filter smoother of (IV-B1) requires each of the forward and reverse passes to be carried out only once, resulting in $\mathcal{O}(|Z_{\text{ev}}|)$ operations. Consequently, the two-filter smoother exhibits the most favourable complexity, although stability issues may preclude its use in certain applications.

V. Example Problem

The mathematical framework of the previous sections is illustrated by way of an example that considers the problem of verifying information sources used in tracking a vessel. Estimates of the vessel's kinematic states were computed by fusing position observations from two ground radar stations, received Automatic Identification System (AIS) messages, and satellite radar (Fig. 1). Without loss of generality, it was assumed that the sources produced measurements simultaneously at a rate of one per second. In this example, measurements and vessel motion were assumed to be linear Gaussian, allowing the application of Kalman filtering and smoothing. The filter and

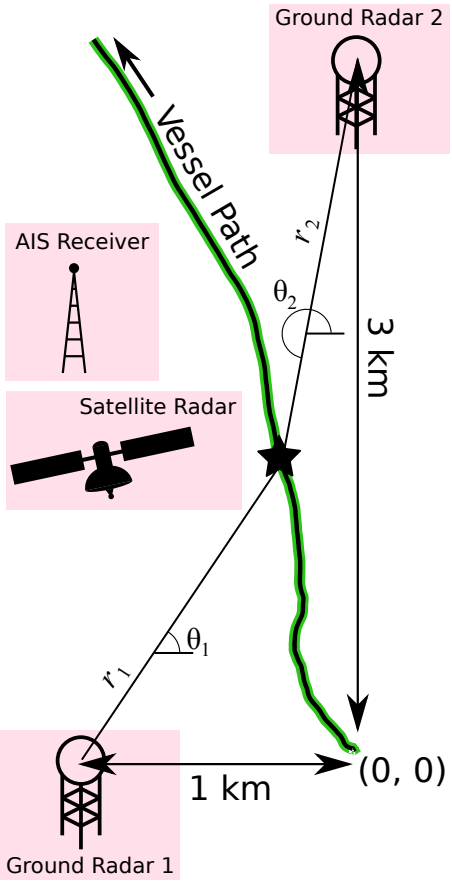

Fig. 1. Multisensor tracking of a vessel showing sensors and their locations. AIS and space-based radar coverage is assumed to be complete over the course of the simulations.

simulated target motion and measurement were implemented using common process and measurement noise covariances given in Table I, where the rotation matrices are defined as

$$\mathbf{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{38}$$

for the pair of θ defined in Fig. 1. To maintain a Gaussian framework, the crescent-shaped measurement likelihood functions ordinarily associated with ground radar were replaced with Cartesian-coordinate approximations, a step that is justified by assuming a small variance to the angular component in the likelihood function.

Two simulations were performed. In the first, the range components of the second ground-radar measurements were scaled by 0.9 (Fig. 2), thereby simulating a malfunction in the time base. In the second, the range components were left unchanged, thus modeling the absence of any faults. Each simulation was carried out for 100 time steps, yielding $n = 200$ degrees of freedom in Z_{ev}. Throughout, the detection and false-alarm probabilities were 1 and 0, respectively, and the transformed samples were generated from the set of measurements using the two-filter smoother of §IV-B2 to reduce the dimensionality of the joint probability density functions.

TABLE I. PARAMETERS OF THE EXAMPLE PROBLEM.

Parameter	Value
Process Noise Cov.	$\begin{bmatrix} 0.25\mathbf{I}\ \mathrm{s}^4 & 0.5\mathbf{I}\ \mathrm{s}^3 \\ 0.5\mathbf{I}\ \mathrm{s}^3 & \mathbf{I}\ \mathrm{s}^2 \end{bmatrix}\ \mathrm{m}^2/\mathrm{s}^4$
Process Noise Mean	$\begin{bmatrix} \mathbf{0}\ \mathrm{m} \\ \mathbf{0}\ \mathrm{m/s} \end{bmatrix}$
Ground Radar 1 Cov. (Pos.)	$\mathbf{R}_{\theta_1}^{-1} \begin{bmatrix} 10^{-2} r_1 \mathrm{m}^{-1} & 0 \\ 0 & 10^{-4} r_1^2 \mathrm{m}^{-2} \end{bmatrix} \mathbf{R}_{\theta_1}\ \mathrm{m}^2$
Ground Radar 1 Mean (Pos.)	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}\ \mathrm{m}$
Ground Radar 2 Cov. (Pos.)	$\mathbf{R}_{\theta_2}^{-1} \begin{bmatrix} 10^{-2} r_2 \mathrm{m}^{-1} & 0 \\ 0 & 10^{-4} r_2^2 \mathrm{m}^{-2} \end{bmatrix} \mathbf{R}_{\theta_2}\ \mathrm{m}^2$
Ground Radar 2 Mean (Pos.)	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}\ \mathrm{m}$
Satellite Radar Cov. (Pos.)	$\begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix}\ \mathrm{m}^2$
Satellite Radar Mean (Pos.)	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}\ \mathrm{m}$
AIS Cov. (Pos.)	$\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}\ \mathrm{m}^2$
AIS Mean (Pos.)	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}\ \mathrm{m}$

Note that $\mathbf{I}$, r_1, and r_2 are the 2×2 identity matrix, distance (in meters) between target and ground radar 1, and distance (in meters) between target and ground radar 2, respectively. The first and second coordinates in centre matrices defining the ground radar covariances are radial and polar, respectively. The 2×2 rotation matrices $\mathbf{R}_{\theta_1}$ and $\mathbf{R}_{\theta_2}$ are defined with respect to the angles given in Fig. 1.

Hypothesis tests were carried out for the information-source partition that assigned ground radar 2 to Z_{ev} and the collection of ground radar 1, space-based radar, and AIS to Z_{tr}. The resultant Kolmogorov–Smirnov test statistics are given in Table II, which clearly illustrates the rejection of H_0 in the defective-radar run, and the non-rejection of H_0 in the no-fault run. In both cases, the results of the test hold for all significance levels between at least 0.2 and 0.1, as given in [17]. The Kolmogorov–Smirnov statistic was also computed under the complete set of single source exclusions. As shown in Fig. 3, exclusion of ground radar 2 yields a statistic of ~ 0.5, which is well below the critical value given by 0.2 significance level. Conversely, exclusion of other sources (and hence inclusion of ground radar 2) yields a statistic of ~ 50, clearly illustrating a lack of consistency that flags ground radar 2 as problematic.

TABLE II. COMPUTED KOLMOGOROV–SMIRNOV STATISTIC.

Simulation Run	Scaled Kolmogorov–Smirnov Statistic
Non-Defective Ground Radar 2	0.74
Defective Ground Radar 2	31.89

Scaled Kolmogorov–Smirnov test statistic ($\sqrt{n}D$) for ground radar 2 in both the defective-radar and no-fault runs. Ground radar 1, space-based radar, and AIS were designated as trusted. Note that for a significance level of 0.01 and 0.2, the asymptotic critical values are $\frac{1.63}{\sqrt{n}}$ and $\frac{1.07}{\sqrt{n}}$, respectively [17].

VI. CONCLUSION

This work studied the problem of assessing information-source performance in statistically-characterized fusion systems. The derived mathematical framework allows information from a collection of suspect sources to be evaluated for consistency by exploiting statistical dependencies with information obtained from trusted sources. The approach is based on a Rosenblatt transformation that casts input data into a sample of independent, identically-distributed observations of the uniform distribution. The transformed sample is then used to compute the Kolmogorov–Smirnov statistic, which provides a quantitative measure for goodness of fit and serves as the basis for an algorithm that detects the presence of suspect information. The mathematical framework is applied to an example that illustrates the identification of a malfunctioning radar system by using the data available from a group of sensors. This work is expected to be of significant value in scenarios that lack ground truth but require validation of received information.

VII. APPENDIX

Definition VII.1 (Rosenblatt Transformation [18], [19])**.** Given an n-dimensional continuous

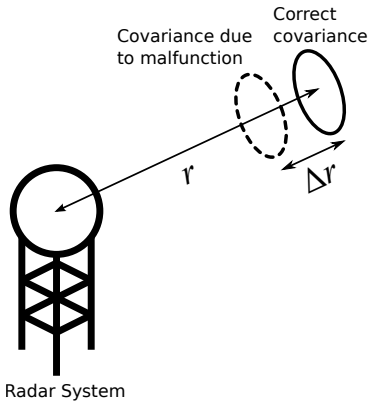

Fig. 2. Error in range due to defective time base. The radial component of the measurement exhibits an error proportional to the distance between the target and radar system (Δr). The variance in bearing is assumed to be small enough to allow the covariance to be approximated in Cartesian coordinates.

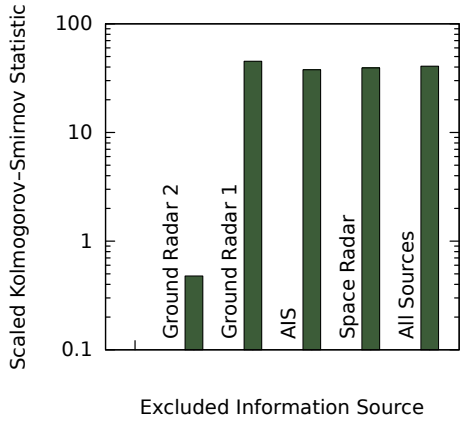

Fig. 3. Scaled Kolmogorov–Smirnov statistic in the defective-radar run as a function of source exclusion (given by $\sqrt{n}D$, where n is the number of degrees of freedom in Z_{ev}). Note the two-order-of-magnitude difference between ground radar 2 and the other sensors.

random vector $Z = \{Z^1, \ldots, Z^n\}$ and its marginal cumulative distribution functions $F_{Z^1}(z^1)$, $F_{Z^2|Z^1}(z^2|z^1), \ldots, F_{Z^n|Z^{n-1},\ldots,Z^1}(z^n|z^{n-1}, \ldots, z^1)$, a new random vector $Y = R(X)$ over $\mathcal{U}^n(0,1)$ may be defined as

$$\begin{aligned} Y^1 &= R(X)^1 &&= F_{Z^1}\left(z^1\right) \\ Y^2 &= R(X)^2 &&= F_{Z^2|Z^1}\left(z^2|z^1\right) \\ &\vdots \\ Y^n &= R(X)^n &&= F_{Z^n|Z^{n-1},\ldots,Z^1}\left(z^n|z^{n-1}, \ldots, z^1\right) \end{aligned} \tag{39}$$

where Y^i and Y^j are statistically independent for $i \neq j$, and $R(\cdot)$ is the Rosenblatt transform. □

Definition VII.2 (Kolmogorov–Smirnov Test [17], [20])**.** A univariate sample $\{x^0, x^1, \ldots, x^n\}$ yields an empirical cumulative distribution function

$$\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} H\left(x - x^i\right) \tag{40}$$

where $H(\cdot)$ is the Heaviside step function. A statistic D may be defined as

$$D = \sup\left|\hat{F}(x) - F(x)\right| \tag{41}$$

for some analytically-defined cumulative distribution function $F(x)$. A goodness-of-fit test may be constructed by defining a null hypothesis asserting that the sample elements are drawn from a population with the same distribution as $F(x)$. The null hypothesis can be rejected at the significance level α if

$$D > K_{\alpha,n} \tag{42}$$

for the critical value $K_{\alpha,n}$, which may be found numerically or from a reference table (cf. [17], [20]). □

ACKNOWLEDGEMENT

The author would like to thank Tim Hammond (DRDC, Atlantic Research Centre) for constructive comments during the preparation of this manuscript.

REFERENCES

[1] S. Wellington, J. Atkinson, and R. Sion, "Sensor Validation and Fusion using the Nadaraya–Watson Statistical Estimator," in *Proceedings of the Fifth International Conference on Information Fusion*, vol. 1. IEEE, 2002, pp. 321–326.

[2] K. Goebel and W. Yan, "Correcting Sensor Drift and Intermittency Faults with Data Fusion and Automated Learning," *IEEE Systems Journal*, vol. 2, no. 2, pp. 189–197, 2008.

[3] P. Aarabi, "Localization-Based Sensor Validation using the Kullback–Leibler Divergence," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 34, no. 2, pp. 1007–1016, 2004.

[4] M. Kumar, D. P. Garg, and R. A. Zachery, "A Generalized Approach for Inconsistency Detection in Data Fusion from Multiple Sensors," in *American Control Conference*. IEEE, 2006, pp. 6–pp.

[5] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, 3rd ed. New York: Springer Science+Business Media, LLC, 2005.

[6] R. Wilcox, *Introduction to Robust Estimation and Hypothesis Testing*, 3rd ed. Oxford, United Kingdom: Academic Press, 2012.

[7] D. M. Mason and J. H. Schuenemeyer, "A Modified Kolmogorov–Smirnov Test Sensitive to Tail Alternatives," *The Annals of Statistics*, vol. 11, no. 3, pp. 933–946, 1983.

[8] Z. Drezner, O. Turel, and D. Zerom, "A Modified Kolmogorov–Smirnov Test for Normality," *Communications in Statistics-Simulation and Computation*, vol. 39, no. 4, pp. 693–704, 2010.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, New Jersey: John Wiley & Sons, 2006.

[10] L. D. Stone, C. A. Barlow, and T. L. Corwin, *Bayesian Multiple Target Tracking*. Norwood, MA: Artech House, 1999.

[11] R. P. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Norwood, MA: Artech House, 2007.

[12] I. R. Goodman, R. P. Mahler, and H. T. Nguyen, *Mathematics of Data Fusion*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1997.

[13] F. Zhang, *The Schur Complement and its Applications*. New York: Springer, 2005.

[14] S. Puntanen, G. P. Styan, and J. Isotalo, *Matrix Tricks for Linear Statistical Models: Our Personal Top Twenty*. Berlin Heidelberg: Springer–Verlag, 2011.

[15] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge, United Kingdom: Cambridge University Press, 2013.

[16] G. Kitagawa, "Non-Gaussian State-Space Modeling of Nonstationary Time Series," *Journal of the American Statistical Association*, vol. 82, no. 400, pp. 1032–1041, 1987.

[17] W. H. Beyer, *CRC Handbook of Tables for Probability and Statistics*. Boca Ratton, Florida: CRC Press, 1966.

[18] M. Rosenblatt, "Remarks on a Multivariate Transformation," *The Annals of Mathematical Statistics*, pp. 470–472, 1952.

[19] H. C. Thode, *Testing for Normality*. New York: CRC Press, 2002.

[20] P. Bickel and K. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*. San Francisco: Holden–Day, 1977.